

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

27  
27  
1/4  
17

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
20 March 2003 (20.03.2003)

PCT

(10) International Publication Number  
WO 03/022052 A1

(51) International Patent Classification<sup>7</sup>: A01N 43/04, C07H 21/02, 21/04, C12N 15/00

(21) International Application Number: PCT/US02/29215

(22) International Filing Date:  
13 September 2002 (13.09.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/322,031 13 September 2001 (13.09.2001) US  
60/347,782 9 January 2002 (09.01.2002) US  
60/389,592 18 June 2002 (18.06.2002) US  
60/406,436 27 August 2002 (27.08.2002) US

(71) Applicant (for all designated States except US): CALIFORNIA INSTITUTE OF TECHNOLOGY [US/US]; 1200 E. California Boulevard, Pasadena, CA 91125 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): BALTIMORE, David [US/US]; 415 South Hill Avenue, Pasadena, CA 91106 (US). QIN, Xiao-Feng [CN/US]; 255 S. Wilson Ave., Apt. 10, Pasadena, CA91106 (US). LOIS-CABALLE, Carlos [ES/US]; 1821 1/2 Wayne Avenue, South Pasadena, CA 91030 (US).

(74) Agent: ALTMAN, Daniel, E.; KNOBBE, MARTENS, OLSON AND BEAR, LLP, 2040 Main Street, Fourteenth Floor, Irvine, CA 92614 (US).

(81) Designated States (national): AE, AG, AL, AM, AT (utility model), AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ (utility model), CZ, DE (utility model), DE, DK (utility model), DK, DM, DZ, EC, EE (utility model), EE, ES, FI (utility model), FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK (utility model), SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:

— of inventorship (Rule 4.17(iv)) for US only

Published:

— with international search report  
— before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/022052 A1

(54) Title: METHOD FOR EXPRESSION OF SMALL RNA MOLECULES WITHIN A CELL

(57) Abstract: The invention provides methods and compositions for the expression of small RNA molecules within a cell using a lentiviral vector. The methods can be used to express double stranded RNA complexes. Small interfering RNA (siRNA) can be expressed using the methods of the invention within a cell, which are capable of down regulating the expression of a target gene through RNA interference. A variety of cells can be treated according to the methods of the invention including embryos, embryogenic stem cells, allowing for the generation of transgenic animals or animals constituted partly by the transduced cells that have a specific gene or a group of genes down regulated.

## METHOD FOR EXPRESSION OF SMALL RNA MOLECULES WITHIN A CELL

### Government Support

[0001] This invention was made with government support under Grant Number GM39458 awarded by the National Institutes of Health. The United States Government has certain rights in the invention.

### Background of the Invention

#### Field of the Invention

[0002] The present invention relates generally to methods for altering gene expression in a cell or an animal using viral constructs engineered to deliver an RNA molecule. In a more specific aspect, a viral construct is used to deliver double-stranded RNA molecules that can be used to down-regulate or modulate gene expression.

#### Description of the Related Art

[0003] RNA interference (RNAi) or silencing is a recently discovered phenomenon (A. Fire *et al.*, *Nature* **391**, 806 (1998); C.E. Rocheleau *et al.* *Cell* **90**, 707 (1997)). Small interfering RNAs ("siRNAs") are double-stranded RNA molecules that inhibit the expression of a gene with which they share homology. siRNAs have been used as a tool to down regulate the expression of specific genes in a variety of cultured cells as well as in invertebrate animals. A number of such approaches have been reviewed recently (P.D. Zamore *Science* **296**, 1265 (2002)); however, such approaches have limitations. For example, no technique prior to the invention described herein allows for the generation of transgenic mammal having a specific gene down regulated through RNA interference. Similarly, there is a need for more robust methods for the introduction of small RNA molecules with regulatory function. The invention provided herein addresses these and other limitations in the field of RNA mediated gene regulation.

### Summary of the Invention

[0004] The invention relates generally to methods to express within a cell an RNA molecule or molecules. These methods can be used with a wide variety of cell types.

RNA molecules can be expressed within a cell for a variety of purposes. For example, and without limitation, RNA molecules can serve as markers within a cell, can act as antisense oligonucleotides or ribozymes for regulating gene expression, and can serve to down regulate genes through RNA interference.

[0005] In one aspect, the invention provides retroviral constructs for the expression of an RNA molecule or molecules within a cell. The constructs preferably comprise a nucleic acid having the R and U5 sequences from a 5' lentiviral long terminal repeat (LTR), a self-inactivating lentiviral 3' LTR, and a RNA Polymerase III (pol III) promoter. The retroviral constructs preferably comprise an RNA coding region operably linked to the RNA Polymerase III promoter. The RNA coding region preferably comprises a DNA sequence that can serve as a template for the expression of a desired RNA molecule.

[0006] The RNA coding region can be immediately followed by a pol III terminator sequence which directs the accurate and efficient termination of RNA synthesis by pol III. The pol III terminator sequences generally comprise 4 or more consecutive T residues. In a preferred embodiment, a cluster of 5 consecutive Ts is used as the terminator by which pol III transcription is stopped at second or third T of the DNA template. As a result, only 2 to 3 U residues are added to the 3' end of the RNA that is synthesized from the RNA coding region.

[0007] A variety of pol III promoters can be used with the invention, including for example, the promoter fragments derived from H1 RNA genes or U6 sn RNA genes of human or mouse origin or from any other species. In addition, pol III promoters can be modified/engineered to incorporate other desirable properties such as to be inducible by small chemical molecules either ubiquitously or in a tissue-specific manner, for example, one activated with tetracycline or IPTG (lacI system).

[0008] The pol III promoter, RNA template region and pol III terminator together may comprise an "RNA cassette" or "RNA expression cassette." If the RNA is a small inhibitory RNA (siRNA), the expression cassette may be termed an "siRNA expression cassette."

[0009] In one embodiment, the RNA coding region encodes a self-complementary RNA molecule having a sense region, an antisense region and a loop region. Such an RNA molecule when expressed desirably forms a "hairpin" structure. The loop region is generally between about 2 and about 10 nucleotides in length. In a preferred

embodiment, the loop region is from about 6 and about 9 nucleotides in length. In one such embodiment of the invention, the sense region and the antisense region are between about 15 and about 30 nucleotides in length.

[0010] In one embodiment, the RNA coding region is operably linked downstream to an RNA Polymerase III promoter such that the RNA coding sequence can be precisely expressed without any extra non-coding nucleotides present at 5' end. In this way an RNA sequence can be expressed that is identical to a target sequence at the 5' end. The synthesis of the RNA coding region is ended at the terminator site. In one preferred embodiment the terminator consists of five consecutive T residues.

[0011] In another aspect of the invention, the retroviral vector can comprise multiple RNA coding regions. In one embodiment, the retroviral construct comprises a first RNA pol III promoter, a first coding region encoding a first RNA molecule operably linked to the first RNA pol III promoter, a second RNA pol III promoter and a second RNA coding region operably linked to the second RNA pol III promoter. Preferably, the second RNA coding region encodes an RNA molecule that is substantially complementary to the RNA molecule encoded by the first RNA coding region, such that the two RNA molecules can form a double-stranded structure when expressed. The methods of invention also include multiple RNA coding regions that encode hairpin-like self-complementary RNA molecules or other non-hairpin molecules.

[0012] In yet another embodiment of the invention, the retroviral construct comprises a first RNA pol III promoter operably linked to a first RNA coding region, and a second RNA pol III promoter operably linked to the same first RNA coding region in the opposite direction, such that expression of the RNA coding region from the first RNA pol III promoter results in a synthesis of a first RNA molecule as the sense strand and expression of the RNA coding region from the second RNA pol III promoter results in synthesis of a second RNA molecule as an antisense strand that is substantially complementary to the first RNA molecule. In one such embodiment, both RNA Polymerase III promoters are separated from the RNA coding region by termination sequences, preferably termination sequences having five consecutive T residues.

[0013] According to one embodiment of the invention, the 5' LTR sequences in the retroviral construct are derived from HIV. The retroviral construct may also comprise a

woodchuck hepatitis virus enhancer element sequence and/or a tRNA amber suppressor sequence.

[0014] In another embodiment of the invention, the self-inactivating 3' LTR is a U3 element with a deletion of its enhancer sequence. In yet another embodiment, the self-inactivating 3' LTR is a modified HIV 3' LTR.

[0015] The recombinant retroviral construct can be pseudotyped, for example with the vesicular stomatitis virus envelope glycoprotein.

[0016] In another aspect of the invention, expression of the RNA coding region results in the down regulation of a target gene. Preferably the target gene comprises a sequence that is at least about 90% identical with the RNA coding region, more preferably at least about 95% identical, and even more preferably at least about 99% identical.

[0017] According to a further aspect of the invention, the viral construct also comprises a nucleotide sequence encoding a gene of interest. The gene of interest is preferably operably linked to a Polymerase II promoter. Such a construct also can contain, for example, an enhancer sequence operably linked with the Polymerase II promoter.

[0018] A variety of Polymerase II promoters can be used with the invention, including for example, the CMV promoter. The RNA Polymerase II promoter that is chosen can be a ubiquitous promoter, capable of driving expression in most tissues, for example, the human Ubiquitin-C promoter, CMV  $\beta$ -actin promoter or PGK promoter. In other embodiments the RNA Polymerase II promoter is a tissue-specific promoter.

[0019] In one embodiment, the gene of interest is a marker or reporter gene, that can be used to verify that the vector was successfully transfected or transduced and its sequences expressed. In one such embodiment, the gene of interest is a fluorescent reporter gene, for example, the Green Fluorescent Protein. In yet another embodiment, the gene of interest is a drug resistant gene which can be used to select the cells that are successfully transduced. For example, the drug resistant gene can be the zeocin resistant gene (zeo). The gene of interest also can be a hybrid of a drug resistant gene and a fluorescent reporter gene, such as a zeo/gfp fusion. In another embodiment, the gene of interest encodes a protein factor that can regulate the transcription activity of inducible pol III promoters. In one of such embodiment, the gene of interest is tetR (repressor for tet operon) which regulates tetracycline responsive pol III promoters.

[0020] It is another aspect of the invention to provide methods for expressing an RNA molecule or molecules within a cell. According to the invention, a packaging cell line is transfected with a retroviral construct of the invention, recombinant retroviral particles are recovered from the packaging cell line; and a target cell is infected with the recombinant retrovirus particles.

[0021] In one embodiment of the invention, the target cell is an embryonic cell. An embryonic cell may be, for example, a single cell embryo or embryonic cells from within an early-stage embryo. In another embodiment of the invention, the target cell is an embryogenic stem cell. When the target cell is an embryonic cell, in one embodiment the embryonic cell is infected by injecting the recombinant retrovirus between the zona pellucida and the cell membrane of the embryonic cell. In another embodiment, the embryonic cell is infected by removing the zona pellucida and incubating the cell in solution containing the recombinant retrovirus. In such an embodiment, the zona pellucida can be removed, for example, by enzymatic digestion.

[0022] When the target cell is an embryonic cell or an embryogenic stem cell, the cell may be transplanted in a pseudopregnant female to generate a transgenic animal.

[0023] The methods of the invention also can be used with a variety of primary ex vivo normal or diseased cells or cells adapted in various tissue culture conditions from human, mouse and other vertebrates, including, without limitation, stem or precursor cells for the hematopoietic system, central nerve system cells, cells with regenerative capacities from a variety of other tissues and organs, dendritic cells and other developing and mature myeloid and lymphoid cells, and cancer cells derived from different cell lineages.

[0024] In a particular embodiment, the target cell is an embryonic cell of a bird within an egg. The embryonic cell of a bird is preferably infected by contacting the embryonic blastodisc of the bird egg with retroviral particles.

[0025] In yet another embodiment, the target cell is a fish egg. The fish egg is preferably infected by delivering the retroviral particles to the space between the chorion and the cell membrane of the fish egg.

#### Brief Description of the Drawings

[0026] Figure 1A shows a schematic diagram of a retroviral vector carrying an expression cassette for RNA expression, termed "RNA cassette" and a "Marker Gene" or gene of interest. The RNA expression cassette can be embedded at any permissible sites of

the retroviral construct either as single copy or multiple tandem copies. In addition, although not indicated in the figure, more than one RNA expression cassette may be present in the retroviral construct. Figure 1B shows a similar construct in which the RNA expression cassettes flank a marker gene.

[0027] Figure 2 shows a schematic view of an RNA expression cassette comprising an RNA Polymerase III promoter 100 linked to an RNA coding region 110-130 and a terminator sequence 140. The RNA coding region comprises a sense region 110, a loop region 120, and an antisense region 130,

[0028] Figure 3 shows a schematic view of an RNA expression cassette having an RNA Polymerase III promoter 100 linked to a first RNA coding region 110 and a first terminator sequence 140 and a second RNA polymerase III promoter 105 linked to a second RNA coding region 115 and a second terminator 145.

[0029] Figure 4 shows a schematic view of an RNA expression cassette having a first RNA Polymerase III promoter 100 linked to an RNA coding region 110 and a first terminator sequence 145. The expression cassette has a second RNA polymerase III promoter 105 linked to the RNA coding region 115, the same sequence as 110 in reverse, and a second terminator 140.

[0030] Figure 5. Schematic illustration of a lacZ siRNA encoding lentiviral vector. 5'LTR: an HIV based lentiviral vector 5' LTR; F: an HIV Flap element; pol III: a human H1-RNA pol III promoter (-240 to -8); siRNA: a lacZ specific small hairpin RNA coding region and its structure and detailed sequence are illustrated below. UbiC: an internal human ubiquitinC promoter; GFP: a GFP marker gene driven by UbiC promoter. W: a woodchuck RNA regulatory element. 3'LTR: an HIV based self inactivating lentiviral 3' LTR.

[0031] Figure 6. A lacZ specific siRNA encoded by a lentiviral vector can efficiently inhibit the expression of lacZ reporter gene in virus transduced mammalian cells. MEF: mouse embryonic fibroblasts; HEK293: human embryonic kidney cells. Both of the test cell lines harbor lacZ and firefly luciferase reporter genes, and the expression levels of the reporter genes can be measured by chemiluminescent assays. Ctrl: the ratio of lacZ activity versus Luc activity of the uninfected parental cells, which was arbitrarily set to 1. Transduced: the specific inhibition of lacZ expression calculated as the reduction of lacZ to Luc ratio.

[0032] Figure 7. Transgenic animals that express a lacZ specific siRNA molecule encoded by a lentiviral vector can successfully suppress the expression of the ubiquitous lacZ reporter gene in a ROSA26<sup>+/−</sup> background. ROSA1-6: the lacZ activities in the limb tissues of six E17.5 ROSA26<sup>+/−</sup> embryos which served as positive controls. The difference in lacZ activity between individual ROSA26<sup>+/−</sup> embryos may result from variable protein extraction efficiency. TG1-4: the lacZ activities in the limb tissues of four E17.5 transgenic embryos expressing a lentiviral vector-encoded lacZ siRNA molecule in ROSA<sup>+/−</sup> background. WT1-6: lacZ activity in the limb tissues of six E17.5 C57Bl/6 wildtype embryos, included as the negative control. The background levels of endogenous beta-galactosidase activity are general below 1,000 LU/ug, thus the columns are not visible.

[0033] Figure 8 shows a schematic illustration of a Tet-inducible lacZ siRNA lentiviral vector. A Tet repressor gene (TetR; SEQ ID NO: 5) is under the control of a human UbiquitinC promoter and its expression can be monitored by the downstream GFP marker coupled by IRES element (internal ribosomal entry site). The anti-lacZ siRNA cassette is driven by a Tet-inducible pol III promoter derived from human U6-promoter (-328 to +1) containing a single TetR binding site (TetO1) between the PSE and TATA box (SEQ ID NO: 4). In the absence of tetracycline, TetR binds to the promoter and its expression is repressed. Upon the addition of tetracycline, TetR is moved from the promoter and transcription starts.

[0034] Figure 9 shows the results of an experiment that demonstrated that a Tet-inducible siRNA expression cassette can regulate gene expression in response to Doxycycline treatment. lacZ and luciferase double expressing HEK293 cells (293Z+Luc) were transduced with a lentiviral vector carrying a Tet-inducible lacZ-siRNA cassette and a Tet repressor under the control of a UbiquitinC promoter (Figure 8). The transduced cells were treated with 10 ug/ml Doxycycline (Plus Dox) for 48hr or without the Doxycycline treatment as a control (No Dox). LacZ and luciferase activities were measured as described in the previous figures. The relative suppression activity is calculated as the ratio of lacZ versus luciferase and No Dox control was arbitrarily set to 1.

#### Detailed Description of the Preferred Embodiment

[0035] The inventors have previously identified a method for introducing a transgene of interest into a cell or animal. This technique is described in co-pending U.S.

provisional patent application number 60/322,031 filed on 9/13/2001 and co-pending U.S. provisional patent application number 60/347,782 filed on 1/9/2002, the entire contents of which are incorporated herein by reference.

[0036] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Any methods, devices and materials similar or equivalent to those described herein can be used in the practice of this invention.

[0037] By "transgene" is meant any nucleotide sequence, particularly a DNA sequence, that is integrated into one or more chromosomes of a host cell by human intervention, such as by the methods of the present invention. In one embodiment, a transgene is an "RNA coding region." In another embodiment the transgene comprises a "gene of interest." In other embodiments the transgene can be a nucleotide sequence, preferably a DNA sequence, that is used to mark the chromosome where it has integrated. In this situation, the transgene does not have to comprise a gene that encodes a protein that can be expressed.

[0038] A "gene of interest" is a nucleic acid sequence that encodes a protein or other molecule that is desirable for integration in a host cell. In one embodiment, the gene of interest encodes a protein or other molecule the expression of which is desired in the host cell. In this embodiment, the gene of interest is generally operatively linked to other sequences that are useful for obtaining the desired expression of the gene of interest, such as transcriptional regulatory sequences.

[0039] A "functional relationship" and "operably linked" mean, without limitation, that the gene is in the correct location and orientation with respect to the promoter and/or enhancer that expression of the gene will be affected when the promoter and/or enhancer is contacted with the appropriate molecules.

[0040] An "RNA coding region" is a nucleic acid that can serve as a template for the synthesis of an RNA molecule, such as an siRNA. Preferably, the RNA coding region is a DNA sequence.

[0041] A "small interfering RNA" or "siRNA" is a double-stranded RNA molecule that is capable of inhibiting the expression of a gene with which it shares homology. In one embodiment the siRNA may be a "hairpin" or stem-loop RNA molecule, comprising a sense region, a loop region and an antisense region complementary to the

sense region. In other embodiments the siRNA comprises two distinct RNA molecules that are non-covalently associated to form a duplex.

[0042] The term "transgenic" is used herein to describe the property of harboring a transgene. For instance, a "transgenic organism" is any animal, including mammals, fish, birds and amphibians, in which one or more of the cells of the animal contain nucleic acid introduced by way of human intervention, such as by the methods described herein. In a transgenic animal that comprises a transgene that encodes a gene of interest, the transgene typically causes the cell to express or overexpress a recombinant protein. However, according to the methods of the invention, expression of an RNA coding region can be used to down regulate the expression of a particular gene through antisense or RNA interference mechanisms.

[0043] The terms "founder," "founder animal" and "founder line" refer to those animals that are mature products of the embryos or oocytes to which the transgene was added, i.e. those animals that grew from the embryos or oocytes into which DNA was inserted.

[0044] The terms "progeny" and "progeny of the transgenic animal" refer to any and all offspring of every generation subsequent to the originally transformed animal.

[0045] The term "animal" is used in its broadest sense and refers to all animals including mammals, birds, fish, reptiles and amphibians.

[0046] The term "mammal" refers to all members of the class Mammalia and includes any animal classified as a mammal, including humans, domestic and farm animals, and zoo, sports or pet animals, such as mouse, rabbit, pig, sheep, goat, cattle and higher primates.

[0047] The term "oocyte" refers to a female gamete cell and includes primary oocytes, secondary oocytes and mature, unfertilized ovum. As used herein, the term "egg" when used in reference to a mammalian egg, means an oocyte surrounded by a zona pellucida. The term "zygote" refers to a fertilized ovum. The term "embryo" broadly refers to an animal in the early stages of development.

[0048] "Perivitelline space" refers to the space located between the zona pellucida and the cell membrane of a mammalian egg or embryonic cell.

[0049] "Target cell" or "host cell" means a cell that is to be transformed using the methods and compositions of the invention.

[0050] "Lentivirus" refers to a genus of retroviruses that are capable of infecting dividing and non-dividing cells. Several examples of lentiviruses include HIV (human immunodeficiency virus; including HIV type 1, and HIV type 2), the etiologic agent of the human acquired immunodeficiency syndrome (AIDS); visna-maedi, which causes encephalitis (visna) or pneumonia (maedi) in sheep, the caprine arthritis-encephalitis virus, which causes immune deficiency, arthritis, and encephalopathy in goats; equine infectious anemia virus, which causes autoimmune hemolytic anemia, and encephalopathy in horses; feline immunodeficiency virus (FIV), which causes immune deficiency in cats; bovine immune deficiency virus (BIV), which causes lymphadenopathy, lymphocytosis, and possibly central nervous system infection in cattle; and simian immunodeficiency virus (SIV), which cause immune deficiency and encephalopathy in sub-human primates.

[0051] A lentiviral genome is generally organized into a 5' long terminal repeat (LTR), the gag gene, the pol gene, the env gene, the accessory genes (nef, vif, vpr, vpu) and a 3' LTR. The viral LTR is divided into three regions called U3, R and U5. The U3 region contains the enhancer and promoter elements. The U5 region contains the polyadenylation signals. The R (repeat) region separates the U3 and U5 regions and transcribed sequences of the R region appear at both the 5' and 3' ends of the viral RNA. See, for example, "RNA Viruses: A Practical Approach" (Alan J. Cann, Ed., Oxford University Press, (2000)), O Narayan and Clements J. Gen. Virology 70:1617-1639 (1989), Fields et al. Fundamental Virology Raven Press. (1990), Miyoshi H, Blomer U, Takahashi M, Gage FH, Verma IM. *J Virol.* 72(10):8150-7 (1998), and U.S. Patent No. 6,013,516.

[0052] Lentiviral vectors are known in the art, including several that have been used to transfect hematopoietic stem cells. Such vectors can be found, for example, in the following publications, which are incorporated herein by reference: Evans JT et al. *Hum Gene Ther* 1999;10:1479-1489; Case SS, Price MA, Jordan CT et al. *Proc Natl Acad Sci USA* 1999;96:2988-2993; Uchida N, Sutton RE, Friera AM et al. *Proc Natl Acad Sci USA* 1998;95:11939-11944; Miyoshi H, Smith KA, Mosier DE et al. *Science* 1999;283:682-686; Sutton RE, Wu HT, Rigg R et al. Human immunodeficiency virus type 1 vectors efficiently transduce human hematopoietic stem cells. *J Virol* 1998;72:5781-5788.

[0053] "Virion," "viral particle" and "retroviral particle" are used herein to refer to a single virus comprising an RNA genome, *pol* gene derived proteins, *gag* gene derived proteins and a lipid bilayer displaying an envelope (glyco)protein. The RNA genome is

usually a recombinant RNA genome and thus may contain an RNA sequence that is exogenous to the native viral genome. The RNA genome may also comprise a defective endogenous viral sequence.

**[0054]** A "pseudotyped" retrovirus is a retroviral particle having an envelope protein that is from a virus other than the virus from which the RNA genome is derived. The envelope protein may be from a different retrovirus or from a non-retroviral virus. A preferred envelope protein is the vesicular stomatitis virus G (VSV G) protein. However, to eliminate the possibility of human infection, viruses can alternatively be pseudotyped with ecotropic envelope protein that limit infection to a specific species, such as mice or birds. For example, in one embodiment, a mutant ecotropic envelope protein is used, such as the ecotropic envelope protein 4.17 (Powell et al. *Nature Biotechnology* 18(12):1279-1282 (2000)).

**[0055]** The term "provirus" is used to refer to a duplex DNA sequence present in a eukaryotic chromosome that corresponds to the genome of an RNA retrovirus. The provirus may be transmitted from one cell generation to the next without causing lysis or destruction of the host cell.

**[0056]** A "self-inactivating 3' LTR" is a 3' long terminal repeat (LTR) that contains a mutation, substitution or deletion that prevents the LTR sequences from driving expression of a downstream gene. A copy of the U3 region from the 3' LTR acts as a template for the generation of both LTR's in the integrated provirus. Thus, when the 3' LTR with an inactivating deletion or mutation integrates as the 5' LTR of the provirus, no transcription from the 5' LTR is possible. This eliminates competition between the viral enhancer/promoter and any internal enhancer/promoter. Self-inactivating 3' LTRs are described, for example, in Zufferey et al. *J. Virol.* 72:9873-9880 (1998), Miyoshi et al. *J. Virol.* 72:8150-8157 and Iwakuma et al. *Virology* 261:120-132 (1999).

**[0057]** The term "RNA interference or silencing" is broadly defined and includes all posttranscriptional and transcriptional mechanisms of RNA mediated inhibition of gene expression, such as those described in (P.D. Zamore *Science* 296, 1265 (2002)).

**[0058]** In one aspect of the invention, a recombinant retrovirus is used to deliver a transgene comprising an RNA coding region of interest to a target cell. Preferably the target cell is a mammalian cell. The cell may be a primary cell, or may be a cultured cell, for example an without limitation an HEK, CHO, COS, MEF, 293 cell. In one embodiment

the target cell is an oocyte or an embryonic cell, more preferably a one-cell embryo. The RNA coding region and any associated genetic elements are thus integrated into the genome of the target cell as a provirus. When the target cell is an embryo, the cell may then be allowed to develop into a transgenic animal by methods well known in the art.

**[0059]** The recombinant retrovirus used to deliver the RNA coding region is preferably a modified lentivirus, and thus is able to infect both dividing and non-dividing cells. The recombinant retrovirus preferably comprises a modified lentiviral genome that includes the transgene. Further, the modified lentiviral genome preferably lacks endogenous genes for proteins required for viral replication, thus preventing undesired replication, such as replication in a resulting transgenic animal. The required proteins are preferably provided in *trans* in the packaging cell line during production of the recombinant retrovirus, as described below.

**[0060]** In another embodiment, the recombinant retrovirus used to deliver the RNA coding region is a modified Moloney virus, for example a Moloney Murine Leukemia Virus. In a further embodiment, the virus is a Murine Stem Cell Virus (Hawley, R. G., et al. (1996) Proc. Natl. Acad. Sci. USA 93:10297-10302; Keller, G., et al. (1998) Blood 92:877-887; Hawley, R. G., et al. (1994) Gene Ther. 1:136-138). The recombinant retrovirus also can be a hybrid virus such as that described in Choi, JK; Hoanga, N; Vilardi, AM; Conrad, P; Emerson, SG; Gewirtz, AM. (2001) Hybrid HIV/MSCV LTR Enhances Transgene Expression of Lentiviral Vectors in Human CD34+ Hematopoietic Cells. *Stem Cells* 19, No. 3, 236-246.

**[0061]** In the preferred embodiment the transgene is incorporated into a viral construct that comprises an intact retroviral 5' LTR and a self-inactivating 3' LTR. The viral construct is preferably introduced into a packaging cell line that packages viral genomic RNA based on the viral construct into viral particles with the desired host specificity. Viral particles are collected and used to infect the host cell. Each of these aspects is described in detail below.

#### The Viral Construct

**[0062]** The viral construct is a nucleotide sequence that comprises sequences necessary for the production of recombinant retrovirus in a packaging cell. In one

embodiment the viral construct additionally comprises genetic elements that allow for the desired expression of an RNA molecule or gene of interest in the host.

[0063] Generation of the viral construct can be accomplished using any suitable genetic engineering techniques well known in the art, including, without limitation, the standard techniques of PCR, oligonucleotide synthesis, restriction endonuclease digestion, ligation, transformation, plasmid purification, and DNA sequencing, for example as described in Sambrook et al. (Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, N.Y. (1989)), Coffin et al. (Retroviruses. Cold Spring Harbor Laboratory Press, N.Y. (1997)) and "RNA Viruses: A Practical Approach" (Alan J. Cann, Ed., Oxford University Press, (2000)).

[0064] The viral construct may incorporate sequences from the genome of any known organism. The sequences may be incorporated in their native form or may be modified in any way. For example, the sequences may comprise insertions, deletions or substitutions. In the preferred embodiment the viral construct comprises sequences from a lentivirus genome, such as the HIV genome or the SIV genome.

[0065] The viral construct preferably comprises sequences from the 5' and 3' LTRs of a lentivirus. More preferably the viral construct comprises the R and U5 sequences from the 5' LTR of a lentivirus and an inactivated or self-inactivating 3' LTR from a lentivirus. The LTR sequences may be LTR sequences from any lentivirus from any species. For example, they may be LTR sequences from HIV, SIV, FIV or BIV. Preferably the LTR sequences are HIV LTR sequences. The virus also can incorporate sequences from MMV or MSCV.

[0066] The viral construct preferably comprises an inactivated or self-inactivating 3' LTR. The 3' LTR may be made self-inactivating by any method known in the art. In the preferred embodiment the U3 element of the 3' LTR contains a deletion of its enhancer sequence, preferably the TATA box, Sp1 and NF-kappa B sites. As a result of the self-inactivating 3' LTR, the provirus that is integrated into the host cell genome will comprise an inactivated 5' LTR.

[0067] Optionally, the U3 sequence from the lentiviral 5' LTR may be replaced with a promoter sequence in the viral construct. This may increase the titer of virus recovered from the packaging cell line. An enhancer sequence may also be included. Any enhancer/promoter combination that increases expression of the viral RNA genome in the

packaging cell line may be used. In the preferred embodiment the CMV enhancer/promoter sequence is used (U.S. Patent No. 5,168,062; Karasuyama et al *J. Exp. Med.* 169:13 (1989).

[0068] The viral construct also comprises a transgene. The transgene, may be any nucleotide sequence, including sequences that serve as markers for the provirus. Preferably the transgene comprises one or more RNA coding regions and/or one or more genes of interest. Schematic diagrams of exemplary retroviral constructs are shown in Figures 1A and 1B.

[0069] In the preferred embodiment the transgene comprises at least one RNA coding region. Preferably the RNA coding region is a DNA sequence that can serve as a template for the expression of a desired RNA molecule in the host cell. In one embodiment, the viral construct comprises two or more RNA coding regions.

[0070] The viral construct also preferably comprises at least one RNA Polymerase III promoter. The RNA Polymerase III promoter is operably linked to the RNA coding region and can also be linked to a termination sequence. In addition, more than one RNA Polymerase III promoter may be incorporated.

[0071] RNA Polymerase III promoters are well known to one of skill in the art. A suitable range of RNA Polymerase III promoters can be found, for example, in Paule and White. *Nucleic Acids Research.*, Vol 28, pp 1283-1298 (2000), which is hereby incorporated by reference in its entirety. The definition of RNA Polymerase III promoters also include any synthetic or engineered DNA fragment that can direct RNA Polymerase III to transcribe a downstream RNA coding sequence. Further, the RNA Polymerase III (Pol III) promoter or promoters used as part of the viral vector can be inducible. Any suitable inducible Pol III promoter can be used with the methods of the invention. Particularly suited Pol III promoters include the tetracycline responsive promoters provided in Ohkawa and Taira *Human Gene Therapy*, Vol. 11, pp 577-585 (2000) and in Meissner et al. *Nucleic Acids Research*, Vol. 29, pp 1672-1682 (2001), which are incorporated herein by reference.

[0072] In one embodiment the transgene comprises a gene of interest that encodes a protein that is desirably expressed in one or more cells of a transgenic animal, for example, a reporter or marker protein. Preferably the gene of interest is located between the 5' LTR and 3' LTR sequences. Further, the gene of interest is preferably in a functional relationship with other genetic elements, for example transcription regulatory sequences such as promoters and/or enhancers, to regulate expression of the gene of interest in a

particular manner once the transgene is incorporated into the host genome. In certain embodiments, the useful transcriptional regulatory sequences are those that are highly regulated with respect to activity, both temporally and spatially.

[0073] Preferably the gene of interest is in a functional relationship with internal Polymerase II promoter/enhancer regulatory sequences. An "internal" promoter/enhancer is one that is located between the 5' LTR and the 3' LTR sequences in the viral construct and is operably linked to the gene that is desirably expressed.

[0074] The Polymerase II promoter/enhancer may be any promoter, enhancer or promoter/enhancer combination known to increase expression of a gene with which it is in a functional relationship.

[0075] The internal promoter/enhancer is preferably selected based on the desired expression pattern of the gene of interest and the specific properties of known promoters/enhancers. Thus, the internal promoter may be a constitutive promoter. Non-limiting examples of constitutive promoters that may be used include the promoter for ubiquitin, CMV (U.S. Patent No. 5,168,062; Karasuyama et al J. Exp. Med. 169:13 (1989),  $\beta$ -actin (Gunning et al. Proc. Natl. Acad. Sci. USA 84:4831-4835 (1987) and pgk (see, for example, U.S. Patent Nos. 4,615,974 and 5,104,795; Adra et al. Gene 60:65-74 (1987), Singer-Sam et al. Gene 32:409-417 (1984) and Dobson et al. Nucleic Acids Res. 10:2635-2637 (1982)). Alternatively, the promoter may be a tissue specific promoter. Several non-limiting examples of tissue specific promoters that may be used include lck (see, for example, Garvin et al. Mol. Cell Biol. 8:3058-3064 (1988) and Takadera et al. Mol. Cell Biol. 9:2173-2180 (1989)), myogenin (Yee et al. Genes and Development 7:1277-1289 (1993), and thy1 (Gundersen et al. Gene 113:207-214 (1992)). In addition, promoters may be selected to allow for inducible expression of the transgene. A number of systems for inducible expression using such a promoter are known in the art, including the tetracycline responsive system and the lac operator-repressor system. It is also contemplated that a combination of promoters may be used to obtain the desired expression of the gene of interest. The skilled artisan will be able to select a promoter based on the desired expression pattern of the gene in the resulting transgenic animal.

[0076] An internal enhancer may also be present in the viral construct to increase expression of the gene of interest. For example the CMV enhancer (Karasuyama et al J. Exp. Med. 169:13 (1989) may be used in combination with the chicken  $\beta$ -actin

promoter (see, e.g., JP 1990005890-A1). Again, one of skill in the art will be able to select the appropriate enhancer based on the desired expression pattern.

[0077] The gene of interest is not limited in any way and includes any gene that the skilled practitioner desires to have integrated and/or expressed in a transgenic animal. For example, the gene of interest may be one that encodes a protein that serves as a marker to identify cells comprising the provirus. In other embodiments the gene of interest encodes a protein that modifies a physical characteristic of the transgenic animal, such as a protein that modifies size, growth, or tissue composition. In another example the gene of interest may encode a protein of commercial value that may be harvested from the transgenic animal.

[0078] In addition, more than one gene of interest may be placed in functional relationship with the internal promoter. For example a gene encoding a marker protein may be placed after the primary gene of interest to allow for identification of cells that are expressing the desired protein. In one embodiment a fluorescent marker protein, preferably green fluorescent protein (GFP), is incorporated into the construct along with the gene of interest. If a second reporter gene is included, an internal ribosomal entry site (IRES) sequence is also preferably included (U.S. Patent No. 4,937,190). The IRES sequence may facilitate the expression of the reporter gene.

[0079] The viral construct may also contain additional genetic elements. The types of elements that may be included in the construct are not limited in any way and will be chosen by the skilled practitioner to achieve a particular result. For example, a signal that facilitates nuclear entry of the viral genome in the target cell may be included. An example of such a signal is the HIV-1 flap signal.

[0080] Further, elements may be included that facilitate the characterization of the provirus integration site in the genome of the animal. For example, a tRNA amber suppressor sequence may be included in the construct.

[0081] In addition, the construct may contain one or more genetic elements designed to enhance expression of the gene of interest. For example, a woodchuck hepatitis virus responsive element (WRE) may be placed into the construct (Zufferey et al. *J. Virol.* 74:3668-3681 (1999); Deglon et al. *Hum. Gene Ther.* 11:179-190 (2000)).

[0082] A chicken  $\beta$ -globin insulator (Chung et al. *Proc. Natl. Acad. Sci. USA* 94:575-580 (1997)) may also be included in the viral construct. This element has been

shown to reduce the chance of silencing the integrated provirus in the transgenic animal due to methylation and heterochromatinization effects. In addition, the insulator may shield the internal enhancer, promoter and exogenous gene from positive or negative positional effects from surrounding DNA at the integration site on the chromosome.

[0083] Any additional genetic elements are preferably inserted 3' of the gene of interest.

[0084] In a specific embodiment, the viral vector comprises: a cytomegalovirus (CMV) enhancer/promoter sequence; the R and U5 sequences from the HIV 5' LTR; the HIV-1 flap signal; an internal enhancer; an internal promoter; a gene of interest; the woodchuck hepatitis virus responsive element; a tRNA amber suppressor sequence; a U3 element with a deletion of its enhancer sequence; the chicken  $\beta$ -globin insulator; and the R and U5 sequences of the 3' HIV LTR.

[0085] The viral construct is preferably cloned into a plasmid that may be transfected into a packaging cell line. The preferred plasmid preferably comprises sequences useful for replication of the plasmid in bacteria.

#### Production of Virus

[0086] Any method known in the art may be used to produce infectious retroviral particles whose genome comprises an RNA copy of the viral construct described above.

[0087] Preferably, the viral construct is introduced into a packaging cell line. The packaging cell line provides the viral proteins that are required in *trans* for the packaging of the viral genomic RNA into viral particles. The packaging cell line may be any cell line that is capable of expressing retroviral proteins. Preferred packaging cell lines include 293 (ATCC CCL X), HeLa (ATCC CCL 2), D17 (ATCC CCL 183), MDCK (ATCC CCL 34), BHK (ATCC CCL-10) and Cf2Th (ATCC CRL 1430). The most preferable cell line is the 293 cell line.

[0088] The packaging cell line may stably express the necessary viral proteins. Such a packaging cell line is described, for example, in U.S. Patent No. 6,218,181. Alternatively a packaging cell line may be transiently transfected with plasmids comprising nucleic acid that encodes the necessary viral proteins.

[0089] In one embodiment a packaging cell line that stably expresses the viral proteins required for packaging the RNA genome is transfected with a plasmid comprising the viral construct described above.

[0090] In another embodiment a packaging cell line that does not stably express the necessary viral proteins is co-transfected with two or more plasmids essentially as described in Yee et al. (*Methods Cell. Biol.* 43A, 99-112 (1994)). One of the plasmids comprises the viral construct comprising the transgene. The other plasmid(s) comprises nucleic acid encoding the proteins necessary to allow the cells to produce functional virus that is able to infect the desired host cell.

[0091] The packaging cell line may not express envelope gene products. In this case the packaging cell line will package the viral genome into particles that lack an envelope protein. As the envelope protein is responsible, in part, for the host range of the viral particles, the viruses are preferably pseudotyped. Thus the packaging cell line is preferably transfected with a plasmid comprising sequences encoding a membrane-associated protein that will permit entry of the virus into a host cell. One of skill in the art will be able to choose the appropriate pseudotype for the host cell that is to be used. For example, in one embodiment the viruses are pseudotyped with the vesicular stomatitis virus envelope glycoprotein (VSVg). In addition to conferring a specific host range this pseudotype may permit the virus to be concentrated to a very high titer. Viruses can alternatively be pseudotyped with ecotropic envelope proteins that limit infection to a specific species, such as mice or birds. For example, in another embodiment, a mutant ecotropic envelope protein is used, such as the ecotropic envelope protein 4.17 (Powell et al. *Nature Biotechnology* 18(12):1279-1282 (2000)).

[0092] In the preferred embodiment a packaging cell line that does not stably express viral proteins is transfected with the viral construct, a second vector comprising the HIV-1 packaging vector with the *env*, *nef*, 5'LTR, 3'LTR and *vpu* sequences deleted, and a third vector encoding an envelope glycoprotein. Preferably the third vector encodes the VSVg envelope glycoprotein.

[0093] In another embodiment of invention, RNA interference activity of the packaging cells is suppressed to improve the production of recombinant virus. This includes, without limitation, the use of cotransfection or stable transfection of constructs expressing siRNA molecules to inhibit Dicer, an RNase III family member of ribonuclease

which is essential for RNA interference (Hammond et al. *Nat. Rev. Genet.* 2:110-119 (2001)).

[0094] The recombinant virus is then preferably purified from the packaging cells, titered and diluted to the desired concentration.

#### Transgenic Animals

[0095] In order to make transgenic animals, an oocyte or one or more embryonic cells are infected with the recombinant virus produced as described above. One of skill in the art will recognize that the method of infection and the treatment of the cell following infection will depend upon the type of animal from which the cell is obtained. For example, mammalian cells are preferably implanted in a pseudopregnant female following infection while for the generation of transgenic birds or fish, the virus is preferably delivered to a laid egg and thus implantation is not required.

[0096] While early methods of making transgenic animals required the cells to be rapidly dividing, there is no such requirement in the methods of the present invention. Thus the cell may be contacted at any point in development. In the preferred embodiment, a zygote is contacted with the recombinant virus.

[0097] The cells to be infected with the virus may be obtained by any method known in the art and appropriate for the specific species in which it is desired to make a transgenic animal. For example, the recovery of fertilized mouse oocytes is described in Hogan et al. (*Manipulating the Mouse Embryo: A Laboratory Manual*. 2<sup>nd</sup> ed. Cold Spring Harbor Laboratory Press, NY (1994)). A method for obtaining fertilized rat oocytes is described in Armstrong et al. (*Biol. Reprod.* 39,511-518 (1998)).

[0098] It is not necessary that the cells be contacted after fertilization. In one embodiment, the virus is delivered to unfertilized ova. Development may then be initialized, for example by *in vitro* fertilization.

#### Delivery of the Virus

[0099] The virus may be delivered to the cell in any way that allows the virus to infect the cell. Preferably the virus is allowed to contact the cell membrane. Two preferred methods of delivering the virus to mammalian cells, injection and direct contact, are described below.

Injection

[0100] In a first embodiment the virus is injected into the perivitelline space between the zona pellucida and the cell membrane of a single-cell zygote. Preferably less than 50 picoliters of viral suspension is injected, more preferably less than 25 picoliters and even more preferably about 10 picoliters.

[0101] The virus is preferably present in a viral suspension and may be injected by any method known in the art. The viral suspension is preferably injected through a hydraulic injector. More preferably a glass micropipette is used to inject the virus. In one embodiment a micropipette is prepared by pulling borosilicate glass capillary on a pipette puller. The tip is preferably opened and beveled to approximately 10  $\mu\text{m}$ . The lentiviral suspension may be loaded into the micropipette from the tip using gentle negative pressure.

[0102] In one embodiment the cell is stabilized with a holding pipette mounted on a micromanipulator, such as by gentle negative pressure against a fire-polished pipette, and a second micromanipulator is used to direct the tip of a micropipette into the space between the zona pellucida and the cell membrane, where the virus is injected.

Direct contact

[0103] In another embodiment the zona pellucida is removed from the cell to produce a denuded embryo and the cell membrane is contacted with the virus. The zona pellucida may be removed by any method known in the art. Preferably it is removed by enzymatic treatment. For example, treatment with pronase may be used to remove the zona pellucida while the cell membrane is kept intact. Alternatively, the cell may be placed in media at pH at which the zona pellucida dissolves while the cell membrane remains intact. For example the cell may be incubated in an acidic Tyrode's solution at room temperature for several minutes. Once the zona pellucida is removed, any method that allows for the virus to contact the cell membrane may be used. Preferably, the cell is incubated in a solution containing the virus. Even more preferably, the solution is media that facilitates survival of the cell.

[0104] In this embodiment, the cells are preferably contacted with the virus in culture plates. The virus may be suspended in media and added to the wells of a multi-well culture plate. The cells may then be plated in the individual wells. The media containing the virus may be added prior to the plating of the cells or after the cells have been plated.

Preferably individual cells are incubated in approximately 10  $\mu$ l of media. However, any amount of media may be used as long as an appropriate concentration of virus in the media is maintained such that infection of the host cell occurs.

**[0105]** The cells are preferably incubated with the virus for a sufficient amount of time to allow the virus to infect the cells. Preferably the cells are incubated with virus for at least 1 hour, more preferably at least 5 hours and even more preferably at least 10 hours.

**[0106]** Both the injection and direct contact embodiments may advantageously be scaled up to allow high throughput transgenesis. Because of the relative simplicity of the injection technique, it is possible to inject many embryos rapidly. For example, it is possible to inject more than 200 fertilized oocytes in less than one hour. With regard to the direct contact embodiment, any number of embryos may be incubated in the viral suspension simultaneously. This may be accomplished, for example, by plating the desired number of single-cell zygotes in multi-well tissue culture plates containing the virus suspended in media appropriate for the survival and growth of the cells.

**[0107]** In both embodiments, any concentration of virus that is sufficient to infect the cell may be used. Preferably the concentration is at least 1 pfu/ $\mu$ l, more preferably at least 10 pfu/ $\mu$ l; even more preferably at least 400 pfu/ $\mu$ l and even more preferably at least  $1 \times 10^4$  pfu/ $\mu$ l.

**[0108]** Following infection with the virus, the cells are preferably implanted in an animal. More preferably cells infected with the virus are implanted in pseudo-pregnant animals of the same species from which the infected cells were obtained. Methods of creating pseudo-pregnancy in animals and implanting embryos are well known in the art and are described, for example, in Hogan et al. (*Manipulating the Mouse Embryo: A Laboratory Manual*. 2<sup>nd</sup> ed. Cold Spring Harbor Laboratory Press, NY (1994)).

**[0109]** In the preferred embodiment early stage embryos (approximately 0 - 2.5 days p.c.) still with an intact zona pellucida are transferred to the oviduct of timed pseudopregnant female (preferably 0.5 days p.c.), while embryos that have reached the blastocyst stage are transferred to the uterus of timed pseudopregnant females (preferably 2.5 days p.c.). Denuded embryos are preferably cultured in vitro until they reach the morula or blastocyst stage (48 to 72 hours in culture), and are then implanted into appropriately timed pseudopregnant females.

[0110] The embryos and resulting animals may be analyzed, for example for integration of the transgene, the number of copies of the transgene that integrated, the location of the integration, the ability to transmit the transgene to progeny and expression of the transgene. Such analysis may be carried out at any time and may be carried out by any methods known in the art. Standard techniques are described, for example, in Hogan et al. (supra).

[0111] The methods of infecting cells disclosed above do not depend upon species-specific characteristics of the cells. As a result, they are readily extended to all mammalian species.

[0112] Initial experiments with mice indicate that of those animals that develop to full term, 80-90% carried at least one copy of the transgene and that, of these, approximately 85% express the gene of interest. Of the transgenic animals about 25% carry only 1 or 2 copies of the transgene. The highest number of proviral insertions observed was about 30. Of the animals that carried only 1 or 2 copies of the transgene, about 80% expressed the gene of interest.

[0113] As discussed above, the modified retrovirus can be pseudotyped to confer upon it a broad host range. One of skill in the art would also be aware of appropriate internal promoters to achieve the desired expression of a gene of interest in a particular animal species. Thus, one of skill in the art will be able to modify the method of infecting cells to create transgenic animals of any species.

[0114] In one embodiment, transgenic birds are created by delivering a modified retrovirus, as described above, to the primordial germ cells of early stage avian embryos. Freshly laid eggs are obtained and placed in a temperature controlled, humidified incubator. Preferably, the embryonic blastodisc in the egg is gradually rotated to lie on top of the yolk. This may be accomplished by any method known in the art, such as by gently rocking the egg regularly, preferably every 15 minutes. Approximately 36 hours later, the modified retrovirus is delivered into the space between the embryonic disk and the perivitelline membrane. Preferably about 50 nL of viral solution is delivered, more preferably about 100 nL of viral solution is delivered, and even more preferably about 200 nL of viral solution is delivered. The viral solution may be delivered by any method known in the art for delivering compositions to the inside of an egg. In the preferred embodiment a window is opened in the shell, the viral solution is injected through the window and the shell window

is closed. The eggs are preferably incubated until hatching. The eggs will hatch after approximately 20 days, depending upon the particular avian species from which they are obtained. Hatched chicks are preferably raised to sexual maturity and mated. The transgenic offspring of the founder animals may be identified by any method known in the art, such as Southern blot, PCR and expression analysis.

**[0115]** In another embodiment, transgenic fish are created by delivering the modified retrovirus, described above, to single-cell fish embryos. Fertilized fish eggs are collected by any method known in the art. The modified retrovirus is then preferably delivered to the space between the chorion and the cell membrane. This may be accomplished, for example, by loading the modified retrovirus in solution into a glass pipette. The pipette may then be used to pierce the chorion membrane and deliver the viral suspension. Preferably about 50 nL of viral solution is delivered, more preferably about 100 nL of viral solution is delivered, and even more preferably about 200 nL of viral solution is delivered. Injected embryos are preferably returned to a temperature-controlled water tank and allowed to mature. At sexual maturity the founder fish are preferably mated and their progeny analyzed for the presence of the transgene by any method known in the art.

**[0116]** As mentioned above, the methods of the present invention will also prove useful in techniques for identifying genes that are involved in specific biological processes, such as gene trap assays and large-scale mutagenesis screens. Such methods are described in the copending provisional patent applications 60/322,031 filed on 9/13/2001 and copending U.S. provisional patent application 60/347,782 filed on 1/9/2002.

#### Down-regulating Gene Expression in a Target Cell

**[0117]** The methods described herein allow the expression of RNA molecules in cells, and are particularly suited to the expression of small RNA molecules, which can not be readily expressed from a Pol II promoter. According to a preferred embodiment of the invention, an RNA molecule is expressed within a cell in order to down-regulate the expression of a target gene. The ability to down-regulate a target gene has many therapeutic and research applications, including identifying the biological functions of particular genes. Using the techniques and compositions of the invention, it will be

possible to knock-down (or down-regulate) the expression of a large number of genes, both in cell culture and in mammalian organisms.

[0118] In preferred embodiments of the invention, an RNA expression cassette comprises a Pol III promoter and an RNA coding region. The RNA coding region preferably encodes an RNA molecule that is capable of down-regulating the expression of a particular gene or genes. The RNA molecule encoded can, for example, be complementary to the sequence of an RNA molecule encoding a gene to be down-regulated. In such an embodiment, the RNA molecule preferably acts through an antisense mechanism.

[0119] A more preferred embodiment involves the expression of a double-stranded RNA complex, or an RNA molecule having a stem-loop or a so-called "hairpin" structure. As used herein, the term "RNA duplex" refers to the double stranded regions of both the RNA complex and the double-stranded region of the hairpin or stem-loop structure.

[0120] Double stranded RNA has been shown to inhibit gene expression of genes having a complementary sequence through a process termed RNA interference or suppression (see, for example, Hammond et al. *Nat. Rev. Genet.* 2:110-119 (2001)).

[0121] According to the invention, an RNA duplex or siRNA corresponding to a region of a gene to be down-regulated is expressed in the cell. The RNA duplex is substantially identical (typically at least about 80% identical, more preferably at least about 90% identical) in sequence to the sequence of the gene targeted for down regulation. siRNA duplexes are described, for example, in Bummelkamp et al. *Science* 296:550-553 (2202), Caplen et al. *Proc. Natl. Acad. Sci. USA* 98:9742-9747 (2001) and Paddison et al. *Genes & Devel.* 16:948-958 (2002).

[0122] The RNA duplex is generally at least about 15 nucleotides in length and is preferably about 15 to about 30 nucleotides in length. However, a significantly longer RNA duplex can be used effectively in some organisms. In a more preferred embodiment, the RNA duplex is between about 19 and 22 nucleotides in length. The RNA duplex is preferably identical to the target nucleotide sequence over this region.

[0123] When the gene to be down regulated is in a family of highly conserved genes, the sequence of the duplex region can be chosen with the aid of sequence comparison to target only the desired gene. On the other hand, if there is sufficient identity among a family of homologous genes within an organism, a duplex region can be designed that would down regulate a plurality of genes simultaneously.

[0124] The duplex RNA can be expressed in a cell from a single retroviral construct. In the preferred embodiment, a single RNA coding region in the construct is a serves as a template for the expression of a self-complementary hairpin RNA, comprising a sense region, a loop region and an antisense region. This embodiment is illustrated in Figure 2, which shows a schematic view of an RNA expression cassette having an RNA Pol III promoter 100 operatively linked to an RNA coding region, having a sense region 110, a loop region 120, an antisense region 130 and a terminator region 140. The sense 110 and antisense 130 regions are each preferably about 15 to about 30 nucleotides in length. The loop region 120 preferably is about 2 to about 15 nucleotides in length, more preferably from about 4 to about 9 nucleotides in length. Following expression the sense and antisense regions form a duplex.

[0125] In another embodiment, the retroviral construct comprises two RNA coding regions. The first coding region is a template for the expression of a first RNA and the second coding region is a template for the expression of a second RNA. Following expression, the first and second RNA's form a duplex. The retroviral construct preferably also comprises a first Pol III promoter operably linked to the first RNA coding region and a second Pol III promoter operably linked to the second RNA coding region. This embodiment is illustrated in Figure 3, which shows a schematic view of an RNA expression cassette having an RNA Polymerase III promoter 100 linked to a first RNA coding region 110 and a first terminator sequence 140 and a second RNA polymerase III promoter 105 linked to a second RNA coding region 115 and a second terminator 145.

[0126] In yet another embodiment of the invention, the retroviral construct comprises a first RNA Pol III promoter operably linked to a first RNA coding region, and a second RNA Pol III promoter operably linked to the same first RNA coding region in the opposite direction, such that expression of the RNA coding region from the first RNA Pol III promoter results in a synthesis of a first RNA molecule as the sense strand and expression of the RNA coding region from the second RNA Pol III promoter results in synthesis of a second RNA molecule as an antisense strand that is substantially complementary to the first RNA molecule. In one such embodiment, both RNA Polymerase III promoters are separated from the RNA coding region by termination sequences, preferably termination sequences having five consecutive T residues. Figure 4 shows a schematic view of an RNA expression cassette having a first RNA Polymerase III

promoter 100 linked to an RNA coding region 110 and a first terminator sequence 145. The expression cassette has a second RNA polymerase III promoter 105 linked to the RNA coding region 115, the same sequence as 110 in reverse, and a second terminator 140.

[0127] In further embodiments an RNA duplex is expressed using two or more retroviral constructs. In one embodiment, a first retroviral construct is used that directs the expression of a first RNA and a second retroviral construct is used that directs expression of a second RNA that is complementary to the first. Following expression the first and second RNAs form a duplex region. It is preferred, however, that the entire duplex region is introduced using retroviral particles derived from a single retroviral construct. As discussed above, several strategies for expressing a duplex RNA from a single viral construct are shown in Figures 2-4.

[0128] The RNA duplexes may be flanked by single stranded regions on one or both sides of the duplex. For example, in the case of the hairpin, the single stranded loop region would connect the duplex region at one end.

[0129] The RNA coding region is generally operatively linked to a terminator sequence. The pol III terminators preferably comprise of stretches of 4 or more thymidine ("T") residues. In a preferred embodiment, a cluster of 5 consecutive Ts is linked immediately downstream of the RNA coding region to serve as the terminator. In such a construct pol III transcription is terminated at the second or third T of the DNA template, and thus only 2 to 3 uridine ("U") residues are added to the 3' end of the coding sequence.

[0130] The sequence of the RNA coding region, and thus the sequence of the RNA duplex, preferably is chosen to be complementary to the sequence of a gene whose expression is to be downregulated in a cell or organism. The degree of down regulation achieved with a given RNA duplex sequence for a given target gene will vary by sequence. One of skill in the art will be able to readily identify an effective sequence. For example, in order to maximize the amount of suppression in a transgenic animal, a number of sequences can be tested for their efficacy in cell culture prior to generating a transgenic animal.

[0131] The methods of the present invention will find great commercial application, for example in biotechnology, medicine and agriculture. For example, in agriculture the described methods may be used to confer disease resistance by expressing in a cell or organism an siRNA that specifically down-regulates the expression of a gene

associated with a pathogen or disease state. In biotechnology, the ability to rapidly develop large numbers of transgenic animals with desired modulation of specific genes will allow for the analysis of gene function and the evaluation of compounds that potentially modulate gene expression, protein function, and are useful in treating a disease or disorder. In particular, by observing the effect of down-regulating specific genes in transgenic animals, the biological function of those genes may be determined. In medicine the methods of the invention may be used to treat patients suffering from particular diseases or disorders, such as HIV, or to confer immunity or resistance to particular pathogens. For example, specific cells may be infected *in vivo* or *ex vivo* with recombinant retrovirus encoding an siRNA that down-regulates the activity of a gene whose activity is associated with a particular disease or disorder.

**[0132]** The following examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

**[0133]** All patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

### Examples

#### Example 1

**[0134]** An lentiviral construct was constructed by insertion of an siRNA expression cassette into the PacI site of HC-FUGW vector (Figure 5; SEQ ID NO: 2). The siRNA was designed to down-regulate expression of the lacZ gene. The HC-FUGW vector comprised a GFP marker gene operably linked to the human Ubiquitin promoter. The GFP marker was useful for tracking transduction events. The vector also comprised an HIV DNA Flap element to improve the virus titers, and the WRE for high level expression of viral genes. The siRNA expression cassette was composed of a pol III promoter and a small hairpin RNA coding region followed by a pol III terminator site. The pol III promoter (SEQ ID NO:3) was derived from the -240 to -9 region of human H1-RNA promoter and

was cloned as an Eco R1 fragment by PCR amplification from HEK293 genomic DNA. The pol III promoter was connected to the downstream RNA coding region by a 7 base pair linker sequence to ensure that the transcription was precisely initiated at the first nucleotide of the RNA coding sequence. The small hairpin RNA coding region comprised a 19 nt sequence corresponding to the 1900-1918 region of the sense strand of the bacterial beta-galactosidase (lacZ) gene coding sequence and the 19 nt perfect reverse complementary sequence separated by a 9 nt loop region. The terminator was comprised of 5 consecutive thymidine residues linked immediately downstream of the RNA coding sequence. The sequence of the hairpin siRNA is shown in SEQ ID NO: 1.

#### Example 2

[0135] Transduction of cultured mammalian cells with retrovirus derived from the retroviral construct described in Example 1 was achieved (Figure 6). The retroviral vector encoding a small hairpin RNA molecule described in Example 1, was used to transfect cultured mammalian cells that express lacZ. A profound decrease in the expression of the lacZ was observed.

[0136] The lacZ siRNA virus was produced by cotransfection of the retroviral vector, a helper virus plasmid and VSVg expression plasmid in HEK293 cells. The virus particles were harvested from the cell culture supernatants and concentrated by ultracentrifugation. The concentrated virus preparations were used to infect either mouse embryonic fibroblasts (MEF) or HEK293 cells which harbor both lacZ and firefly luciferase (Luc) reporter genes. Infection was monitored by the GFP signal which is expressed from the marker gene cassette of the viral vector. Under the conditions of this experiment, >98% of the test cells were GFP+ and thus were successfully transduced. The expression levels of lacZ and Luc reporter genes were measured by chemiluminescent assays using commercially available kits (lacZ assay kit from Roche and Luc from Promega). The lacZ siRNA virus only inhibited the expression of lacZ but not Luc. The specific inhibition was determined by the ratio of lacZ activity to Luc activity. The lacZ/Luc ratio of the uninfected parental cells was arbitrarily set to 1 and the values of the infected cells were calculated accordingly. As shown in Figure 6, transfection with the virus resulted in dramatic reduction in the amount of expression of the lacZ gene in both MEF and HEK293 cells.

[0137] A tet-inducible lacZ siRNA lentiviral vector was also prepared as illustrated in Figure 8. A Tet repressor gene (TetR; SEQ ID NO: 5) was placed the under the control of the human UbiquitinC promoter so that its expression could be monitored by the downstream GFP marker. The anti-lacZ siRNA cassette was driven by a Tet-inducible pol III promoter derived from human U6-promoter (-328 to +1) containing a single TetR binding site (TetO1) between the PSE and TATA box (SEQ ID NO: 4). The TetR coding sequence was PCR amplified from genomic DNA from the TOP10 strain of E. coli adn cloned into a modified version of FUIGW as a Bgl2-EcoR1 fragment. In the absence of tetracycline, TetR binds to the promoter and its expression is repressed. Upon the addition of tetracycline, TetR is moved from the promoter and transcription starts.

[0138] The Tet-inducible siRNA expression cassette was able to regulate gene expression in response to Doxycycline treatment. Virus was prepared from the retroviral construct carrying the Tet-inducible lacZ-siRNA cassette and a Tet repressor under the control of a UbiquitinC promoter and used to transduce HEK293 cells expressing both lacZ and luciferase (293Z+Luc). The transduced cells were treated with 10 ug/ml Doxycycline (Plus Dox) for 48hr or without the Doxycycline treatment as a control (No Dox). LacZ and luciferase activities were measured as described in the previous figures. The relative suppression activity is calculated as the ratio of lacZ versus luciferase and No Dox control was arbitrarily set to 1. As can be seen in Figure 9, in the presence of doxycycline suppression of lacZ activity was significantly enhanced.

### Example 3

[0139] This example demonstrates the generation of transgenic animals that express an siRNA molecule encoded by a lentiviral vector. The expression of the lacZ specific siRNA described in Example 1 resulted in extensive suppression of lacZ activity in ROSA26+/- mice.

[0140] ROSA26+/- mice carry one copy of a ubiquitously expressed lacZ reporter gene. The lacZ siRNA virus preparations described in Example 2 were used for perivitelline injection of ROSA26+/- single cell embryos obtained from hormone primed C57Bl/6 female donors x ROSA26+/+ stud males. The injected single cell embryos were subsequently transferred into the oviduct of timed pseudopregnant female recipients. Embryonic day 15.5 to 17.5 (E15.5-17.5) fetuses were recovered from the surrogate

mothers. Successful transgenesis was scored by positive GFP signal observed with the fetuses under fluorescent microscope. Protein extracts prepared from the limb tissues of the fetuses were used for the LacZ chemiluminescent assay according to the manufacturer's instruction (Roche), and protein concentrations of the tissue extracts were determined by the Bradford assay (BioRad). The lacZ expression levels were expressed as light units (LU) per ug of proteins (LU/ug). The E15.5-17.5 fetuses from the timed mating of C57Bl/6 females x ROSA26<sup>+/+</sup> males and C57Bl/6 females x C57Bl/6 males were served as the positive and negative controls respectively. The results are shown in Figure 7. Animals G1-G4 (those treated derived from embryos infected with the virus comprising the siRNA construct) showed markedly decreased expression of the lacZ gene as compared with untreated control animals.

## WHAT IS CLAIMED IS:

1. A method of expressing an RNA molecule within a cell, the method comprising:
  - transfected a packaging cell line with a retroviral construct;
  - recovering recombinant retrovirus from the packaging cell line; and
  - infecting a target cell with the recombinant retrovirus,

wherein the retroviral construct comprises the R and U5 sequences from a 5' lentiviral long terminal repeat (LTR), a self-inactivating lentiviral 3' LTR, an RNA Polymerase III promoter region and an RNA coding region operably linked to an RNA Polymerase III promoter region.
2. The method of claim 1, wherein the retroviral construct further comprises at least one termination sequence operably linked to the RNA coding region.
3. The method of claim 1, wherein the RNA Polymerase III promoter is inducible.
4. The method of claim 3, wherein the inducible promoter is activated with tetracycline.
5. The method of claim 1, wherein the RNA coding region encodes a self-complementary RNA molecule having a sense region, an antisense region and a loop region.
6. The method of claim 5, wherein the loop region is about 2 to about 10 nucleotides in length.
7. The method of claim 5, wherein the sense region and the antisense region are each between about 15 and about 30 nucleotides in length.
8. The method of claim 1, wherein the retroviral construct comprises a first RNA coding region operably linked to a first RNA Polymerase III promoter and a second RNA coding region operably linked to a second RNA Polymerase III promoter.
9. The method of claim 8, wherein the first RNA coding region encodes a first RNA molecule and the second RNA coding region encodes a second RNA molecule.
10. The method of claim 9, wherein the first RNA molecule and the second RNA molecule are substantially complementary.
11. The method of claim 1, wherein the retroviral construct comprises a first RNA Polymerase III promoter and a second RNA Polymerase III promoter, each operably

linked to the RNA coding region, such that expression of the RNA coding region from the first RNA Polymerase III promoter results in the synthesis of a first RNA molecule and expression of the RNA coding region from the second RNA Polymerase III promoter results in the synthesis of a second RNA molecule substantially complementary to the first RNA molecule.

12. The method of claim 1, wherein expression of the RNA coding region results in the down regulation of a target gene, wherein the target gene comprises a sequence that is at least about 90% identical with the RNA coding region.
13. The method of claim 1 wherein said packaging cell line is a 293 cell line.
14. The method of claim 1 wherein the 5' LTR sequences are from HIV.
15. The method of claim 1, wherein the viral construct comprises the woodchuck hepatitis virus enhancer element sequence.
16. The method of claim 1, wherein the viral construct comprises a tRNA amber suppressor sequence.
17. The method of claim 1 wherein the self-inactivating 3' LTR comprises a U3 element with a deletion of its enhancer sequence.
18. The method of claim 17, wherein the self-inactivating 3' LTR is a modified HIV 3' LTR.
19. The method of claim 1, wherein the recombinant retrovirus is pseudotyped.
20. The method of claim 19, wherein the recombinant retrovirus is pseudotyped with the vesicular stomatitis virus envelope glycoprotein.
21. The method of claim 1, wherein the viral construct further comprises a gene of interest.
22. The method of claim 21, wherein the viral construct has a Polymerase II promoter operably linked to the gene of interest.
23. The method of claim 22, wherein the promoter is a CMV promoter.
24. The method of claim 22, wherein the viral construct additionally comprises an enhancer operably linked to the promoter.
25. The method of claim 25, wherein the enhancer and promoter are CMV sequences.
26. The method of claim 21, wherein the gene of interest is a reporter gene.

27. The method of claim 26, wherein the reporter gene encodes a fluorescent protein.
28. The method of claim 27, wherein the fluorescent protein is green fluorescent protein.
29. The method of claim 20, wherein the polymerase II promoter is a ubiquitous promoter.
30. The method of claim 27, wherein the ubiquitous promoter is selected from the group consisting of the ubiquitin promoter, the CMV  $\beta$ -actin promoter and the pgk promoter.
31. The method of claim 22, wherein the RNA Polymerase II promoter is a tissue specific promoter.
32. The method of claim 31, wherein said tissue specific promoter is selected from the group consisting of the lck promoter, the myogenin promoter and the thy1 promoter.
33. The method of claim 1, wherein the target cell is an embryonic cell.
34. The method of claim 33, wherein the target cell is an embryogenic stem cell.
35. The method of claim 33, further comprising implanting the embryo in a pseudopregnant female.
36. The method of claim 1, wherein infecting a target cell comprises injecting the recombinant retrovirus between the zona pellucida and the cell membrane of a mammalian embryonic cell.
37. The method of claim 1, wherein infecting a target cell comprises removing the zona pellucida from a mammalian embryonic cell and incubating the cell in solution containing the recombinant retrovirus.
38. The method of claim 37, wherein the zona pellucida is removed by enzymatic digestion.
39. The method of claim 1, wherein the target cell is a cultured cell.
40. The method of claim 39, wherein the target cell is a cultured mammalian cell.
41. The method of claim 40, wherein the cultured mammalian cell is selected from the group consisting of CHO, HEK, COS and MEF cells.
42. The method of claim 1, wherein the target cell is an embryonic cell of a bird.

43. The method of claim 41, wherein infecting an embryonic cell of a bird comprises contacting the embryonic blastodisc of the bird egg with the recombinant retrovirus.

44. The method of claim 1, wherein the target cell is a fish egg.

45. The method of claim 42, wherein infecting the fish egg comprises delivering the recombinant retrovirus to the space between the chorion and the cell membrane of the fish egg.

46. A retroviral construct for the expression of an RNA molecule within a cell, the retroviral construct comprising:

a nucleic acid having the R and U5 sequences from a 5' lentiviral long terminal repeat (LTR);

a self-inactivating lentiviral 3' LTR; and

an RNA Polymerase III promoter.

47. The retroviral construct of claim 46, further comprising at least one termination sequence.

48. The retroviral construct of claim 46, wherein the RNA Polymerase III promoter is inducible.

49. The retroviral construct of claim 48, wherein the inducible promoter is activated with tetracycline.

50. The retroviral construct of claim 46, further comprising at least one RNA coding region operably linked to the RNA Polymerase III promoter.

51. The retroviral construct of claim 50, wherein the RNA coding region encodes a self-complementary RNA molecule having a sense region, an antisense region and a loop region.

52. The retroviral construct of claim 51, wherein the loop region is about 2 to about 10 nucleotides in length.

53. The retroviral construct of claim 51, wherein the sense region and the antisense region are between about 15 and about 30 nucleotides in length.

54. The retroviral construct of claim 50, wherein the RNA coding region encodes a first RNA molecule, and the retroviral construct further comprises a second RNA Polymerase III promoter and a second RNA coding region operably linked to the second

RNA Polymerase III promoter, wherein the second RNA coding region encodes a second RNA molecule substantially complementary to the first RNA molecule.

55. The retroviral construct of claim 50, wherein the retroviral construct further comprises a second RNA Polymerase III promoter operably linked to the RNA coding region, such that expression of the RNA coding region from the first RNA Polymerase III promoter results in a synthesis of a first RNA molecule and expression of the RNA coding region from the second RNA Polymerase III promoter results in synthesis of a second RNA molecule substantially complementary to the first RNA molecule.

56. The retroviral construct of claim 50, wherein expression of the RNA coding region results in the down regulation of a target gene.

57. The method of claim 56, wherein the target gene comprises a sequence that is at least about 90% identical with the RNA coding region.

58. The retroviral construct of claim 46, wherein the 5' LTR sequences are from HIV.

59. The retroviral construct of claim 46, wherein the viral construct comprises the woodchuck hepatitis virus enhancer element sequence.

60. The retroviral construct of claim 46, wherein the viral construct comprises a tRNA amber suppressor sequence.

61. The retroviral construct of claim 46, wherein the self-inactivating 3' LTR comprises a U3 element with a deletion of its enhancer sequence.

62. The retroviral construct of claim 46, wherein the self-inactivating 3' LTR is a modified HIV 3' LTR.

63. The retroviral construct of claim 46, wherein the recombinant retrovirus is pseudotyped.

64. The retroviral construct of claim 63, wherein the recombinant retrovirus is pseudotyped with the vesicular stomatitis virus envelope glycoprotein.

65. The retroviral construct of claim 46, wherein the viral construct further comprises a gene of interest.

66. The retroviral construct of claim 65, wherein the viral construct has a Polymerase II promoter operably linked to the gene of interest.

67. The retroviral construct of claim 66, wherein the RNA Polymerase II promoter is a CMV promoter.

68. The retroviral construct of claim 66, wherein the viral construct additionally comprises an enhancer operably linked to the RNA Polymerase II promoter.

69. The retroviral construct of claim 68, wherein the enhancer and the RNA Polymerase II promoter are CMV sequences.

70. The retroviral construct of claim 66, wherein the RNA Polymerase II promoter is a ubiquitous promoter.

71. The retroviral construct of claim 70, wherein the ubiquitous promoter is selected from the group consisting of the ubiquitin promoter, the CMV  $\beta$ -actin promoter and the pgk promoter.

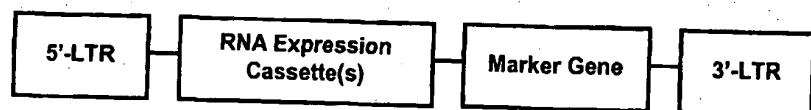
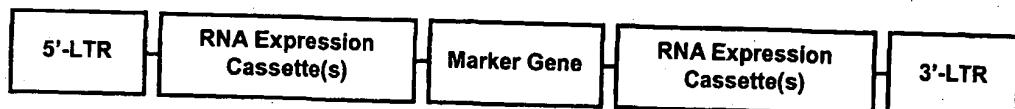
72. The retroviral construct of claim 66, wherein the RNA Polymerase II promoter is a tissue specific promoter.

73. The retroviral construct of claim 72, wherein said tissue specific promoter is selected from the group consisting of the lck promoter, the myogenin promoter and the thy1 promoter.

74. The retroviral construct of claim 65, wherein the gene of interest is a reporter gene.

75. The retroviral construct of claim 74, wherein the reporter gene encodes a fluorescent protein.

76. The retroviral construct of claim 75, wherein the fluorescent protein is green fluorescent protein.

**FIGURE 1A****FIGURE 1B**

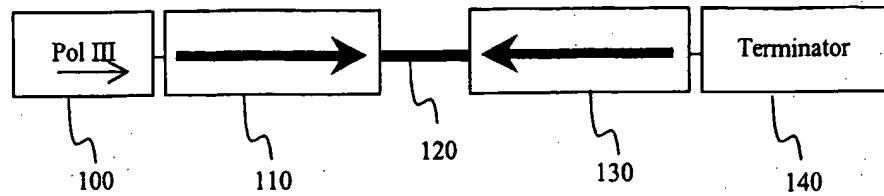
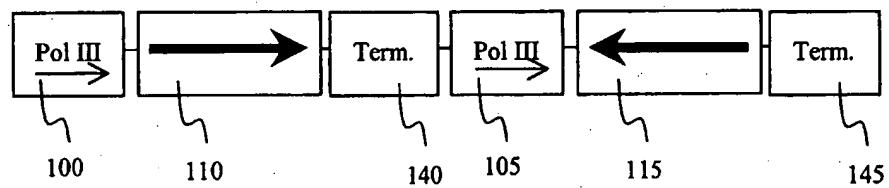
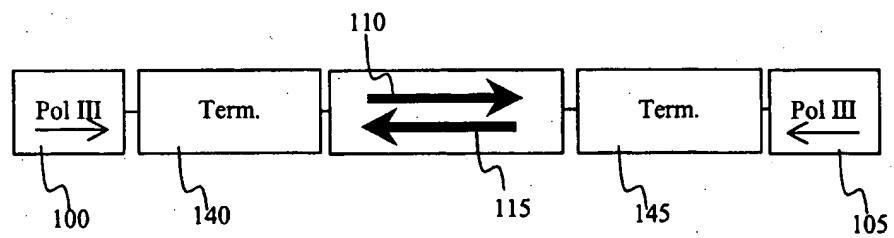
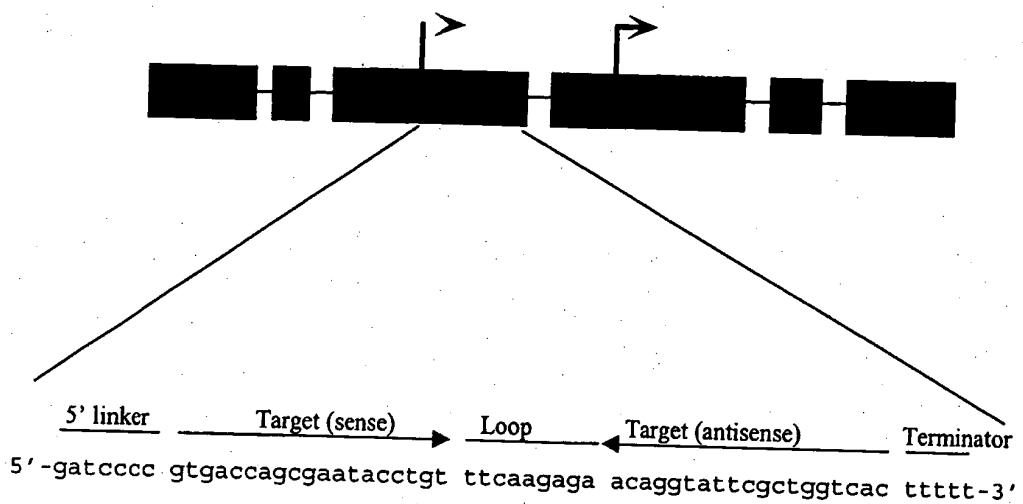
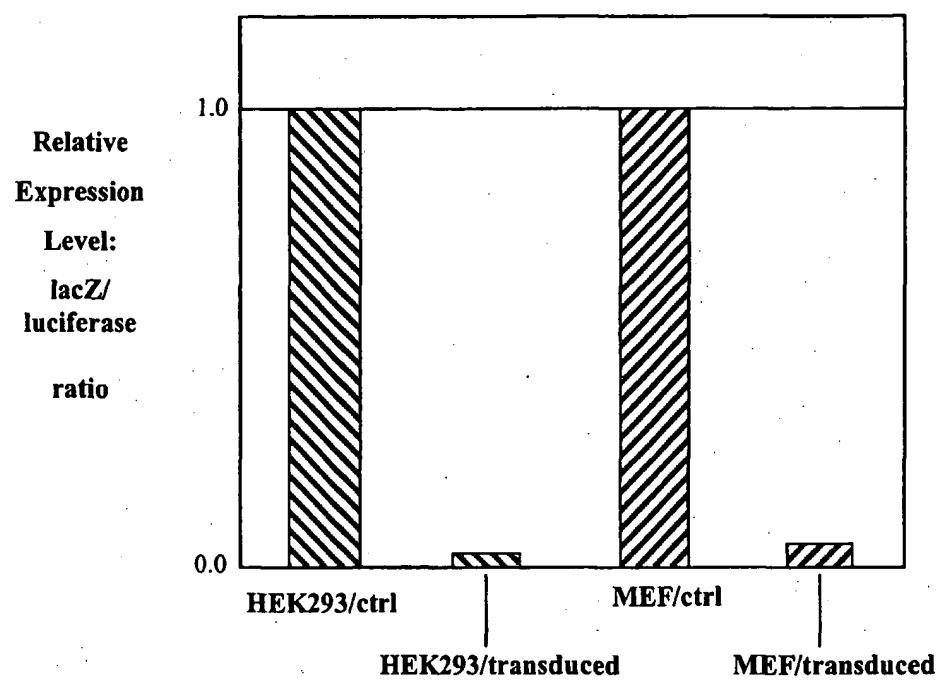
**FIGURE 2****FIGURE 3****FIGURE 4**

FIGURE 5



**FIGURE 6**

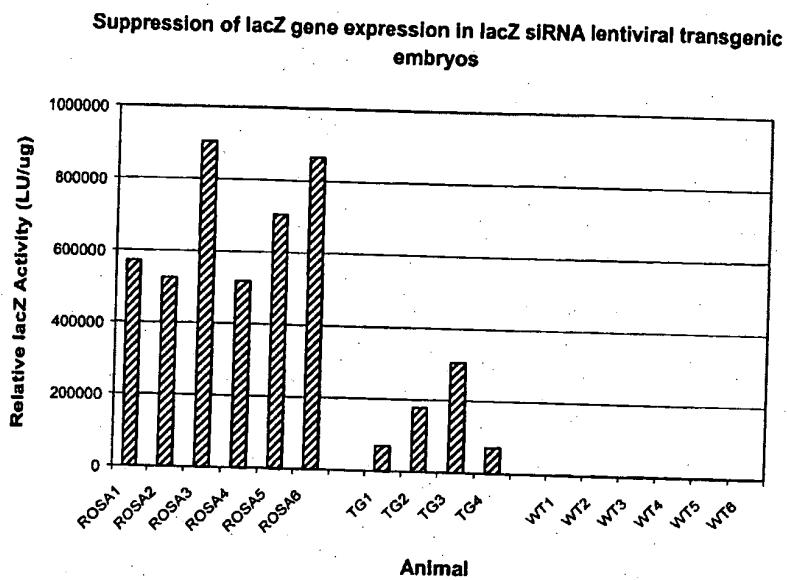
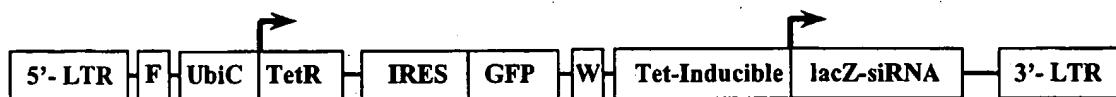
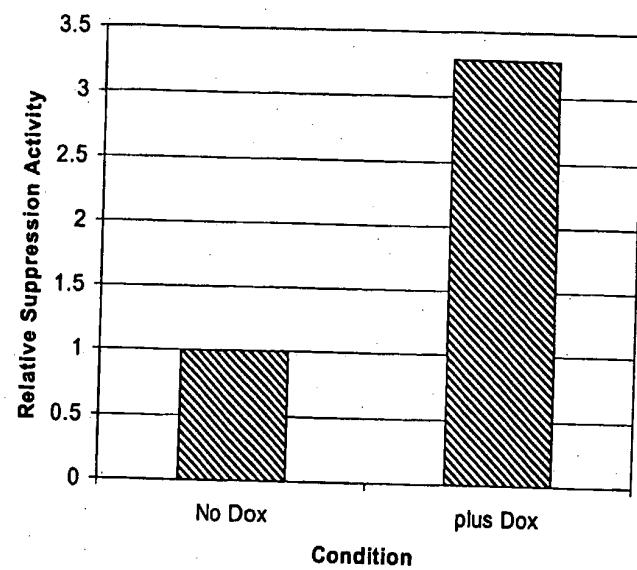
**FIGURE 7**

FIGURE 8



**FIGURE 9****Suppression of Gene Expression by Tet-inducible si-RNA**

## SEQUENCE LISTING

<110> CALIFORNIA INSTITUTE OF TECHNOLOGY  
IN, XIAO-FENG  
LOIS-CABALLE, CARLOS  
BALTIMORE, DAVID

<120> METHOD FOR EXPRESSION OF SMALL RNA  
MOLECULES WITHIN A CELL

<130> CALTE.010VPC

<150> 60/322,031  
<151> 2001-09-13

<150> 60/347,782  
<151> 2002-01-09

<150> 60/389,592  
<151> 2002-06-18

<150> Not yet available (Attorney Docket No.: CALTE.011PR)  
<151> 2002-08-27

<160> 5

<170> FastSEQ for Windows Version 4.0

<210> 1  
<211> 59  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> This represents a siRNA cassette comprising  
bacterial sequence and synthetic linker, loop and  
terminator sequences.

<400> 1  
gatccccgtg accagcgaat acctgtttca agagaacagg tattcgctgg tcacttttt 59

<210> 2  
<211> 9941  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> This sequence represents a lentiviral vector  
comprising a human immunodeficiency virus flap  
sequence, a green fluorescent protein variant  
sequence, a human ubiquitin promoter sequence and  
a woodchuck hepatitis regulator element sequence.

<400> 2  
gtcgacggat cgggagatct cccgatcccc tatggtgac tctcagtaca atctgctctg 60



cagggtgggg ccacctgccc gtaggtgtgc ggtaggctt tctccgtcgc aggacgcagg 3540  
gttcgggcct aggtaggct ctccctgaatc gacaggcgcc ggacctctgg tgaggggagg 3600  
gataagttag gctcgatgtt ctttggtegg ttttatgtac ctatcttctt aagttagctga 3660  
agctccggtt ttgaactatg cgctcggggt tggcgagtgt gttttgtgaa gttttttagg 3720  
cacctttga aatgtaatca ttgggttcaa tatgtaaattt tcagtgtagt actagtaaag 3780  
cttctgcagg tcgactctag aaaattgttcc gctaaattctt ggccgtttt ggctttttg 3840  
ttagacagga tccccgggtt cccgtcgccca ccatggtagt caagggcgag gagctgttca 3900  
ccgggggtgtt gcccattctt gtcgagctgg acggcgacgt aaacggccac aagttcagcg 3960  
tgtccggcga gggcgagggtt gatgccaccc gaccaagct gaccctgaag ttcatctgca 4020  
ccaccggcaa gctgcccgtt ccctggccca ccctcgtagt caccctgacc tacggcggtc 4080  
agtgcctcag ccgcgttccccc gaccacatga agcagcacga ctcttcaag tccggccatgc 4140  
ccgaaggtt cgtccaggag cgccacatct tcttcaagga cgacggcaac tacaagaccc 4200  
gcccggaggtt gaagttcgag ggcgacaccc tggtaaccg catcgagctg aagggcatcg 4260  
acttcaagga ggacggcaac atcctggggc acaagctgga gtacaactac aacagccaca 4320  
acgtctatata catggccgac aagcagaaga acggcatcaa ggtgaactt aagatccggc 4380  
acaacatcga ggacggcagc gtgcagctc ccggaccaacta ccacgagaac acccccatcg 4440  
gacgacggccc cgtgctgtgtt cccgacaacc actacctgag caccctgtcc gcccgtgagca 4500  
aagaccccaa cgagaagcgc gatcacaatgg tccctgtggta gttcgtagcc gccggccgg 4560  
tcactctcggtt catggacggat ctgtacaagt aaagcgcccg cgactctaga attcgatatac 4620  
aagcttatacg ataatcaacc tctggattac aaaattgtt aagatttgac tggtattttt 4680  
aactatgttgc tcccttttac gctatgtggta tacgctgtt taatgcctt gtatcatgt 4740  
attgcttccc gtatggctttt cattttctcc tccctgtata aatctgtttt gctgtctttt 4800  
tatgaggagt tggccgttgcgtt tgcaggccaa cgtggcggtt tgcactgtt gtttgcgtac 4860  
gcaaccccaa ctgtggggggg cattggccacc acctgtcagc tcccttccgg gactttcgct 4920  
ttccccctcc ctattggccac ggcggaaactc atgcggccctt gccttggccg ctgctggaca 4980  
ggggctcgcc tggggccac tgcataattcc tgggtgttgcgtt cggggaaatc atcgcttcc 5040  
ccttggctgc tcgcctgtgtt tgcacccgtt attctggcgccg ggacgtccctt ctgcgtacgtc 5100  
ccttcggccc tcaatccagc ggaccccttcc tcccgccggc tgctgcccgc tcttcggcc 5160  
cttcccggtc ttcgccttcg ccctcagacg agtccgtatcc ccccttggcc cgcctccccc 5220  
catcgatacc gtgcacccgtt agacccgtt aaacatggag caatcacaag tagcaataca 5280  
gcagctacca atgtgttgcgtt tgcctggccaa gaagcacaag aggaggagga ggtgggtttt 5340  
ccagtcacac ctcaggttcc ttaagacca atgacttaca aggcagctgtt agatcttagc 5400  
cacttttaa aaaaaagggg gggactggaa gggctaaatcc actcccaacg aagacaagat 5460  
atcccttgcgtt tggatgttcc ccacacacaa ggctacttcc ctgatggca gaactacaca 5520  
ccagggccag ggatcagata tccactgacc tttggatgttgcgtt gtcataagctt agtaccagg 5580  
gagcaagaga aggttgcgttcc agccaatgaa ggagagaaca cccgcttgcgtt acaccctgt 5640  
agcctgcgttcc ggttgcgttcc cccggagaga gaagtattag agtggaggtt tgacagccgc 5700  
ctagcatttc atcacatggc ccggagacgtt catccggact gtactgggtc tctctggta 5760  
gaccagatctt gggccctggaa gctctctggc taacttagggaa accccactgtt taagcctcaa 5820  
taaagcttgc ttgtgtgttgcgtt tcaagtagtgc tggccgttgc tgggtgtgaa ctctggtaac 5880  
tagagatccc tcagaccctt ttgtgtgttgcgtt tggaaaatctt ctacgaggccgc ccgtttaaac 5940  
ccgctgtatca gcctcgactg tgccttctgtt tggccagcca tctgttgcgtt gcccctccccc 6000  
cgtgccttcc ttgaccctggaa aagggtggccac tcccactgtc ctttccataat aaaatgagga 6060  
aattgcatacg cattgtctgaa gttaggtgtca ttctattctgtt ggggggtggggg tggggcagga 6120  
cagcaagggg gaggattgggaa aagacaatag caggcatgtt ggggtgtccgg tgggtcttat 6180  
ggcttctgttgcgttcc gccccttgcgttcc cccagctgggg ctcttaggggg tttcccaacg cgcctgttag 6240  
cgccgcattttt aactgttgcgttcc gttgtgtgttgcgttcc tacgctcgccgtt gtcgttccaa 6300  
cgccctagcg cccgttccctt tcgcgttccctt cccttccctt ctcggccacgt tcgcccggc 6360  
tccccgtccaa gctctaaatc gggggctccc tttaggggttgcgtt ccgtttagtgc ttttacggca 6420  
cctcgacccccc aaaaaacttgcgttcc tttaggggttgcgtt ccgtttagtgc ttttacggca 6480  
gacggttttt ccccttgcgttcc cccggggccgttcc cccggggccgttcc ttttacggca 6540  
aactggaaaca acactcaacc ctatctcggtt ctattttttt gatttataag ggattttggcc 6600  
gatttcggcc tattggttaa aaaaatgttgcgttcc gatttataacaa aaatttaaacg cgaattaatt 6660  
ctgtggaaatg ttttacggca accaggatgttgcgttcc gaaatggccat cggccctgtata 6720  
atgcataatcgttcc ttttacggca accaggatgttgcgttcc gaaatggccat cggccctgtata 6780  
gcaggccagaa gtatgcataatcgttcc aatttagtgcgttcc gaaatggccat cggccctgtata 6840  
actccggccca tccctggccctt aactccggccca agttccggccca attctccggccca ccatgggttgcgttcc 6900

ctaatttttt ttatattatgc agaggccgag gcccgcctcg cctctgagct attccagaag 6960  
 tagtgaggag gcttttttgg aggcttaggc ttttgcaaaa agctcccgag agcttgata 7020  
 tccattttcg gatctgatca gcacgtgttgc acaattaatc atccgcata tatactggca 7080  
 tagtataata cgacaagggtg aggaactaa ccatggccaa gttgaccagt ggcgttccgg 7140  
 tgctcaccgc gcgcgacgtc gceggagcgg tcgagttctg gaccgaccgg ctcgggttct 7200  
 cccgggactt cgtggaggac gacttcgccc gtgtgttcc ggacgacgtg accctgttca 7260  
 tcagcgcgtt ccaggaccag gtggtgcggg acaacacccct ggcctgggtg tgggtgcgg 7320  
 gcctggacga gctgtacgcc gagtggtcgg aggtcgtgtc cacgaacttc cgggacgcct 7380  
 cccggccggc catgaccgag atcggcgagc agccgtgggg gcccggatgtc gcccctgcgc 7440  
 accccggccgg caactgctgtg cacttcgtgg ccgaggagca ggactgacac gtgctacgag 7500  
 atttcgatcc caccggccgc ttctatgaaa ggttgggctt cggaaatcggtt tccgggacg 7560  
 cccggctgtt gatccctccag cgcggggatc tcatgcttgc gttttcgcc cacccttcaact 7620  
 ttttttttgc agttataat gtttacaaat aaagaatcatc acatcacaat ttccacaataa 7680  
 aagattttt ttcaactgtat tctagttgtt gtttgcctaa actcatcaat gtatcttac 7740  
 atgtctgtat accgtcgacc tctagctaga gtttggcgtt atcatggtca tagctgtttc 7800  
 ctgtgtgaaa ttgttataatccg ctcacaatcc cacacaacat acgagccggaa agcataaaagt 7860  
 gtaaaaggctg ggggtctta ttagttagtgc aactcacatt aatttgcgttgcgtcaact 7920  
 cccgttcca gtcggaaac ctgtcggtcc agtgcattatc atgaatcgcc caacgcgcgg 7980  
 ggagaggcggg ttgcgtatt gggcgcttcc cccgttctc gtcactgtac tcgctgcgt 8040  
 cggtcgttgc gtcggcggca gccgtatcgtt ctcacttcaaa ggcgttataa cgttattcca 8100  
 cagaatcagg ggataacgcga gggaaagaaca ttgtgacaaa aggcacgca aaggccagga 8160  
 accgtaaaaaa ggccgcgttgc ctggcggtt tccataggtt cccggccccc gacgagcatc 8220  
 acaaaaaatcg acgtcataatc agaggtggc gaaaccgcac aggactataa agataccagg 8280  
 cgtttccccc tggaaagctcc ctctgtgcgtt ctccctgttcc gaccctggcc ctaccggat 8340  
 acctgtccgc cttttcccttc tggggaaagcg tggcgcttcc tcatagtcac tcgtgttagt 8400  
 atctcaggcc ggttaggtc gttcgcttca agcccgacccg ctgcgcctta tccggtaact 8460  
 actttatcgcc actggcagca gccactgttgc gtgttacaga gtttcttgcgat ttttttttgc 8520  
 gtgttacaga gtttcttgcgat ttttttttgcgat actacggcttca cactagaaga acatgtatcc 8580  
 atctcgtgcgc tctgtgttgcgat ccagttaatcttcc tccggaaatcc tccggaaatcc 8640  
 gcaaaacaaac caccgttgcgat agccgtgggtt ttttttttgcgat ttttttttgcgat 8700  
 gaaaaaaagg atctcaagaa gatcccttgcgat tcccttgcgat tcccttgcgat 8760  
 acgaaaaacttcc acgtttaaggg attttggtca tggattatc aaaaaggatc ttacacccat 8820  
 tccctttaaa taaaaatgcgat ttttttttgcgat caatctaaatc ttttttttgcgat 8880  
 ctgacaggta ccaatgttca atccgttgcgat cacctatctc agcgatctgtt ttttttttgcgat 8940  
 catccatagt tgcctgtactt cccgtcggtt ttttttttgcgat ttttttttgcgat 9000  
 ctggcccccgg tgcgtcaatgcgat tcccttgcgat tcccttgcgat tcccttgcgat 9060  
 caataaaacca gccaggccggaa agggccggac acccacgcgtc accggcttca gatttattcc 9120  
 ccatccagtc tattattgt tgcggggaaatcc tccggaaatcc tccggaaatcc tccggaaatcc 9180  
 tgegcaacgt ttttttttgcgat tcccttgcgat tcccttgcgat tcccttgcgat 9240  
 cttcatttgcgat tcccttgcgat tcccttgcgat tcccttgcgat tcccttgcgat 9300  
 aaaaaggcggt tagcttccgcgat tcccttgcgat tcccttgcgat tcccttgcgat 9360  
 tatcaactcat gtttatggca gcaactgcata ttttttttgcgat tcccttgcgat 9420  
 gcttttctgt gactgggttagt tactcaatcc tcccttgcgat tcccttgcgat 9480  
 cgagttgtcc ttttttttgcgat tcccttgcgat tcccttgcgat tcccttgcgat 9540  
 aagtgtccat cattggaaaaatcc tcccttgcgat tcccttgcgat tcccttgcgat 9600  
 tgagatccat ttttttttgcgat tcccttgcgat tcccttgcgat tcccttgcgat 9660  
 tcaccacgtt ttttttttgcgat tcccttgcgat tcccttgcgat tcccttgcgat 9720  
 gggcgacacgtt gaaatgttgcgat tcccttgcgat tcccttgcgat tcccttgcgat 9780  
 atccgggttgcgat ttttttttgcgat tcccttgcgat tcccttgcgat tcccttgcgat 9840  
 taggggttcc gggcgacatcc tcccttgcgat tcccttgcgat tcccttgcgat 9900  
 gggcgacacgtt ttttttttgcgat tcccttgcgat tcccttgcgat tcccttgcgat 9941

<210>

<211> 233

<212> DNA

<213> *Homo sapiens*

<400> 3

gaattcgaac gctgacgtca tcaacccgct ccaaggaatc gcggggccag tgcactagg 60  
cgggaacacc cagcgcgcgt gcgcctggc aggaagatgg ctgtgaggga caggggagtg 120  
gcgcctgca atatttgcat gtcgctatgt gttctggaa atcaccataa acgtgaaatg 180  
tctttgatt tggaatctt ataagttctg tatgagacca cagatctaag ctt 233

<210> 4  
<211> 355  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> This represents a mutant human sequence having an  
introduced bacterial tetO1 binding site.

<400> 4  
gggaattccc ccagtggaaa gacgcgcagg caaaacgcac cacgtgacgg agcgtgaccg 60  
cgcgcgcgac ccaaggtcg gcaggaagag ggcctattc ccatgattcc ttcatatttg 120  
catatacgat acaaggctgt tagagagata attagaatta atttgactgt aaacacaag 180  
atatttagtac aaaatacgtg acgtagaaaag taataatttc ttgggtagtt tgcagtttt 240  
aaattatgtt taaaaatgga ctatcatatg cttaccgtaa cttgaaagta ctctatcatt 300  
gatagaggtt tataatctgtt ggaaaggacg aaacaccgtg gtcttcaagc ttccg 355

<210> 5  
<211> 634  
<212> DNA  
<213> E. coli

<400> 5  
gctagccacc atgtccagat tagataaaag taaagtgatt aacagcgcac tagagctgct 60  
taatgaggc ggaatcgaag gtttacaac ccgtaaactc gcccagaagc taggtgtaga 120  
gcagccatac ttgtattggc atgtaaaaaa taagcggct ttgtctcgacg ccttagccat 180  
tgagatgtt gataggcacc atactcactt ttgccttta gaaggggaaa gctggcaaga 240  
ttttttacgt aataacgcata aaagttagt atgtgtttta ctaagtcatc gcgatggagc 300  
aaaagtacat tttagtacac ggcctacaga aaaacagtat gaaactctcg aaaatcaatt 360  
agcctttta tgccacaacaag gtttttcaact agagaatgca ttatatgcac tcagcgtgt 420  
ggggcattttt acttttaggtt gcgtattggaa agatcaagag catcaagtgc ctaaagaaga 480  
aaggaaaca cctactactg atagtatgcc gccattatta cgacaagctc tcgaattatt 540  
tgatcaccaa ggtcagacg cagccttctt attcggcattt gaattgtatca tatgcggatt 600  
agaaaaacaa cttaaatgtg aaagtgggtc tttaa 634

**INTERNATIONAL SEARCH REPORT**

International application No.

PCT/US02/29215

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC(7) : A01N 43/04; C07H 21/02, 21/04; C12N 15/00  
US CL : 514/44; 536/23.1; 536/24.5; 435/320.1

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)  
U.S. : 514/44; 536/23.1; 536/24.5; 435/320.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
EAST, STN-Medline Biosis EMBASE Lifesci CAplus

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 6,096,538 A (KINGSMAN et al) 1 August 2000 (01.08.2000), column 2, lines 6-12; column 7, lines 41-43; column 3, lines 57-67, Figure 2B.	1-3, 8, 9, 17, 21, 22, 29, 46-48, 50, 61, 62, 66, 70, 74
Y	US 5,883,081 A (KRAUS et al) 16 March 1999 (16.03.1999), column 29, lines 15-60; column 47, Example 10; column 43, lines 65-67); column 48 lines 49-50; column 47, lines 65-66 and Table 2.	1, 14, 22, 23, 29, 30, 39-41, 19, 20, 46, 63, 64, 66, 67, 70, 71
Y	REISER et al. Development of Multigene and Regulated Lentivirus Vectors. Jour. Virol. November 2000, Vol. 74, No. 22, pages 10589-10599.	1-4, 19-30, 39, 40, 46-50, 63-71, 74-76
Y	IWAKUMA et al. Self-inactivating Lentiviral Vectors with U3 and U5 Modifications. Virology, 1999, Vol. 261, pages 120-132.	1, 17, 18, 46, 61, 62
Y	GATLIN et al. Long-term engraftment of nonobese diabetic/severe combined immunodeficient mice with human CD34+ cells transduced by a self-inactivating human immunodeficiency virus type I vector. Human Gene Therapy, 10 June 2001, Vol. 12, pages 1079-1089.	1, 17-18, 46, 61, 62
Y	US 6,255,071 B1 (BEACH et al) 03 July 2001 (03.07.2001), col. 6, lines 43-47; col. 13, lines 60-65; col. 14, lines 35-36	1, 2, 13, 46

Further documents are listed in the continuation of Box C.  See patent family annex.

• Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"B" earlier application or patent published on or after the international filing date	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O" document referring to an oral disclosure, use, exhibition or other means	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"P" document published prior to the international filing date but later than the priority date claimed	"&"	document member of the same patent family

Date of the actual completion of the international search 15 January 2003 (15.01.2003)	Date of mailing of the international search report 10 FEB 2003
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	Authorized officer Valarie Bertoglio Telephone No. 703-308-1234
Facsimile No. (703)305-3230	

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/29215

## C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	GODWIN et al. Detection of targeted GFP-Hox gene fusions during mouse embryogenesis. PNAS, October 1998, Vol. 95, pages 13042-13047, specifically Abstract and page 13042, col. 2, "Vector Construction".	27,28,75,76
Y	TIMMONS et al. Specific interference by ingested dsRNA. Nature, 29 October 1998, Vol. 395, page 854, see whole document.	11,15
Y	ILVES et al. Retroviral Vectors designed for targeted expression of RNA polymerase III-driven transcripts: a comparative study. Gene, 1996, Vol. 171, pages 203-208, specifically page 203 Abstract and Introduction; pages 207, column 1, lines 14-17.	1,8-11,46,54,55
Y	JUNKER et al. Reduction in replication of the human immunodeficiency virus type I in human T cell lines by polymerase III-driven transcription of chimeric tRNA-antisense RNA genes. Antisense Research and Development, 1994, Vol. 4, pages 165-172, specifically page 166, column 1, lines 22-23.	1,8-11,46,54,55
Y	OGUETA et al. Design and in vitro characterization of a single regulatory module for efficient control of gene expression in both plasmid DNA and a self-inactivating lentiviral vector. Molecular Medicine, 2001, Vol. 7, No. 8, pages 569-579.	1,3,4,46,48,49,61,62
Y,P	BRUMMELKAMP et al. A system for stable expression of short interfering RNAs in mammalian cells. 19 April 2002, Vol. 296, pages 550-553, specifically column 3, lines 2-3 and 12-19.	1,5-11,46,54,55
Y	KAFRI et al. Lentiviral Vectors: Regulated gene expression. Molecular Therapy, June 2000, Vol. 1, No. 6, pages 516-521.	4

